

PATENT ABSTRACTS OF JAPAN

(11)Publication number : 11-075157

(43)Date of publication of application : 16.03.1999

(51)Int.Cl.

H04N 5/92
 G11B 20/10
 G11B 20/12
 H03M 7/30
 H04N 5/928
 H04N 7/24

(21)Application number : 09-234594

(71)Applicant : SONY CORP

(22)Date of filing : 29.08.1997

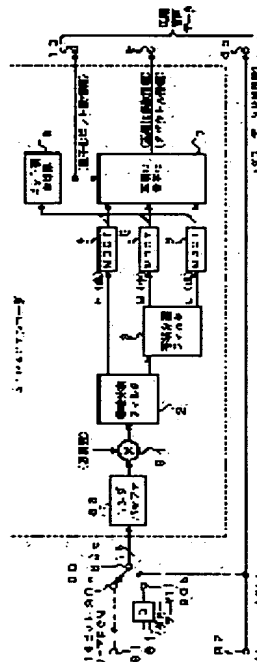
(72)Inventor : SHIGENOBU MASAHIRO
INOUE KOJI

(54) RECORDER FOR VIDEO SIGNAL AND AUDIO SIGNAL

(57)Abstract:

PROBLEM TO BE SOLVED: To evade the variance of amplitude and also to use an ATRAC encoder and an ATRAC decoder which are used currently by recording the video and audio signals on a storage medium synchronously with each other for each GOP after compressing both video and audio signals via the MPEG and ATRAC respectively.

SOLUTION: The video and audio data are compressed by an MPEG 2 and an ATRAC respectively. At the same time, the sound group position information to which the dummy data are added by the dummy data adding means 80 and 81 are stored in the header information and recorded on a recording medium. Then both video and audio data are recorded sequentially on the recording medium for each GOP. The header information, the compressed video data equivalent to a single GOP and then the audio data equivalent to two channels are recorded in a single GOP period. Thus, the amplitude never causes any variance when the sample number of every sound group is multiplied by a fixed window number and then synthesized.



LEGAL STATUS

[Date of request for examination]

[Date of sending the examiner's decision of rejection]

[Kind of final disposal of application other than the examiner's decision of rejection or application converted registration]

[Date of final disposal for application]

[Patent number]

[Date of registration]

This Page Blank (uspto)

【特許請求の範囲】

【請求項1】 映像データをMPEGで符号化し、圧縮映像データを生成する映像符号化手段と、

前記圧縮映像データのGOP単位で一定となるサンプル数が含まれるように音声信号のサンプリング周波数を設定し、該サンプリング周波数で前記音声信号をデジタル化するA/D変換手段と、

該A/D変換手段からの音声データをATRACで符号化すると共に前記圧縮映像データの1GOPに対し一定サウンドグループ数の圧縮音声データを生成する音声符号化手段と、

前記音声符号化手段に供給する音声データのサンプル数を1サウンドグループ当たり一定とするダミーデータを付加するダミーデータ付加手段と、

前記圧縮映像データと前記圧縮音声データとを前記GOP単位で記録する記録手段とを有することを特徴とする映像信号及び音声信号の記録装置。

【請求項2】 請求項1記載の映像信号及び音声信号の記録装置において、

前記音声データの有効サンプル数の情報をヘッダ等別エリアに記録するようにしたことを特徴とする映像信号及び音声信号の記録装置。

【発明の詳細な説明】

【0001】

【発明の属する技術分野】本発明は異なる単位で圧縮が施された映像信号及び音声信号を同期させて記録するようにした映像信号及び音声信号の記録装置に関する。

【0002】

【従来の技術】現在、音声信号に圧縮処理を施し、ディスク又はテープ等の記録媒体に記録し、これを再生する手法として、標準化されたATRAC (Adaptive Transform Acoustic Coding) が知られている。このATRACはMD (Mini Disk)に採用されている。

【0003】このATRACにつき図8、図9を参照して説明する。図8は音声データを圧縮するATRACのエンコーダを示す。

【0004】図8において、例えば、0～22kHzの周波数帯域のオーディオ信号が、例えば、サンプリング周波数が44.1kHzを以ってサンプリングされた後、PCM化されて得られた入力オーディオPCM信号が、入力端子1に供給される。この入力オーディオPCM信号は、例えばいわゆるQMF (クワドラチャ・ミラー・フィルタ) フィルタ等の帯域分割フィルタ2により0～11kHz帯域と11kHz～22kHz帯域とに分割される。更に、0～11kHz帯域の信号は同じくいわゆるQMFフィルタ等の帯域分割フィルタ3により0～5.5kHz帯域と5.5kHz～11kHz帯域とに分割される。

【0005】上述の帯域分割フィルタ2からの11kHz～22kHz帯域の信号は、直交交換回路の一例であ

るMDCT (Modified Discrete Cosine Transform) 回路 (モディファイド・離散コサイン変換手段) (直交変換手段) 4に供給されて、MDCT処理される。帯域分割フィルタ3からの5.5kHz～11kHz帯域の信号はMDCT回路 (モディファイド・離散コサイン変換手段) (直交変換手段) 5に供給されて、MDCT処理される。帯域分割フィルタ3からの0～5.5kHz帯域信号はMDCT回路 (モディファイド・離散コサイン変換手段) (直交変換手段) 6に供給されて、MDCT処理される。なお、各MDCT回路4, 5, 6では、各帯域毎に設けたブロック決定回路により決定されたブロックサイズ (情報圧縮パラメータ) (処理ブロックの長さ) に基づいてMDCT処理がなされる。

【0006】上述したように、入力デジタル信号を複数の周波数帯域に分割する手段としては、例えば、QMFフィルタがあるが、これについては、1976 R.E. Crochiere Digital Coding of Speech In Subbands Bell Syst. Tech. J. Vol. 55, No. 8 1976 に述べられている。又、ICASSP 83, Boston Polyphase Quadrature Filters-A New Subband Coding Technique Joseph H. Rothweilerには、等バンド幅のフィルタ分割方法が述べられている。ここで、上述した直交変換としては、例えば、入力オーディオ信号を所定単位時間 (フレーム) でブロック化し、そのブロック毎に高速フーリエ変換 (FFT)、離散コサイン変換 (DCT)、モディファイドDCT変換 (MDCT) 等を行なうことで、時間軸を周波数軸に変換するようにした直交交換がある。MDCTについてはICASSP 1987 Subband/Transform Coding Using Filter Bank Designs Based On Time Domain Aliasing Cancellation J. P. Princen A. B. Bradley Univ. of Surrey Royal Melbourne Inst. Of Tech. に述べられている。

【0007】この直交交換ブロックサイズの決定は、直交交換ブロックサイズの決定回路で行われ、その決定結果は各MDCT回路4, 5, 6及びビット割当算出回路8に供給されると共に、ブロックのブロックサイズ情報 (処理ブロックの長さの情報) (情報圧縮パラメータ) として出力端子10より出力される。

【0008】各MDCT回路4, 5, 6にてMDCT処理されて得られた周波数軸上のスペクトルデータ、又はMDCT係数データ (時間と周波数に関する2次元ブロック内の信号成分) は、低域はいわゆる臨界帯域 (クリティカルバンド) 毎にまとめられて、中高域はブロックフローティングの有効性を考慮して、臨界帯域幅を細分化して正規化/量子化回路7及びビット割り当て算出回路8に供給される。

【0009】この臨界帯域 (クリティカルバンド) とは、人間の聴覚特性を考慮して分割された周波数帯域であり、ある純音の周波数近傍の同じ強さの狭帯域バンドノイズによって純音がマスクされるときそのノイズの持つ帯域のことである。この臨界帯域 (クリティカルバ

ンド)は、高域ほど帯域幅が広くなっており、上述の0～22kHzの全周波数帯域は例えば25のクリティカルバンドに分割されている。

【0010】ビット割当算出回路8は、上述のブロックサイズ情報及び、スペクトルデータ又はMDCT係数データに基づき、いわゆるマスキング効果等を考慮して、上述の臨界帯域及びブロックフローティングを考慮した各分割帯域毎の、マスキング量、及び、同分割帯域毎のエネルギーあるいはピーク値等を算出し、その結果に基づき、各帯域毎に割当ビット数を求め、正規化／量子化回路7へ供給している。この正規化／量子化回路7では、上述のブロックサイズ情報(情報圧縮パラメータ)(処理ブロックの長さ)、及び、臨界帯域及びブロックフローティングを考慮した各分割帯域毎に割り当てられたビット数に応じて、各スペクトルデータ、又はMDCT係数データを再量子化(正規化して量子化)するようにしている。このようにして符号化されたデータは、図8における出力端子9を介して取り出される。

【0011】斯る図8のエンコードにおいては入力された例えば16ビットの音声データを二段の帯域分割フィルタ2、3によって3つの帯域に分割した後、MDCT4、5、6によってスペクトル信号に変換する。

【0012】このように予め帯域を分割してからスペクトル信号に変換することによって、スペクトル信号が広い帯域に広がることを防いでいる。これによって例えば特にマスキングが効きにくい低域に鋭いトーン性の信号が入力された場合にも高いS/N比を保った圧縮ができるため、充分な音質を保証することができる。

【0013】得られたスペクトル信号は符号化ユニットごとに正規化された後、量子化されるが、これを何ビットで行うか計算するために、量子化されるスペクトル信号そのものを利用できる。以上の処理は512サンプルごとに行われ、符号化されたデータが記録される。

【0014】図9はこのようにして圧縮されたオーディオ信号を復元するデコードを示す。各帯域の量子化されたMDCT係数、即ち、図8における出力端子9の出力信号と等価のデータは、図9における入力端子11に供給され、使用されたブロックサイズ情報(処理ブロックの長さ)(情報圧縮パラメータ)、即ち、図8における出力端子10の出力信号と等価のデータは、図9における入力端子12に供給される。スペクトル復元回路13では適応ビット割当情報を用いてビット割当を解除する。次に逆直交変換(IMDCT)回路14、15、16では周波数軸上の信号が時間軸上の信号に変換される。これらの部分帯域の時間軸上信号は、図9における帯域合成フィルタ(IQMF)回路17、18により、全帯域信号に復号化されて、出力端子19に出力される。

【0015】このデコードではまず記録されているデータからスペクトル信号が復元され、次にこれにエンコー

ダで使われたMDCTの逆変換が施されることによって各帯域の時系列信号が得られ、最後に二段の帯域合成フィルタ17、18によって例えば16ビットのオーディオ信号が合成される。

【0016】また映像信号と音声信号とに圧縮処理を施し、ディスク又はテープ等の記録媒体に記録し、これを再生する手法として、標準化されたMPEG2(Moving Picture Experts Group 2)が知られている。このMPEG2はDVD(Digital Video Disk)やデジタル放送等に採用されている。

【0017】このMPEG2では映像信号の圧縮が15フレーム(約0.5秒)単位で行なわれるものが主流となりつつある。この圧縮の単位をGOP(Group of Pictures)と呼ぶ。NTSC方式において、このGOPの時間は、正確には1001/2000秒と規定されており、この単位で映像信号と音声信号との同期を取るためには、音声信号のサンプリング周波数とGOP周波数(2000/1001 Hz)がある関係、例えば整数比の関係になければいけない。

【0018】具体的には、MDやCD等で用いられるサンプリング周波数は、44.1kHzであるがこの周波数では、1GOP(1001/2000秒)の中のサンプル数が22,072.05個と小数になり、映像信号と音声信号との同期を取るのは困難であった。即ちこれら標準化された圧縮の手法において、相互間で利用することが困難であった。

【0019】斯る点に鑑み先にATRAC及びMPEG2の異なる単位で圧縮した音声信号及び映像信号を同期させて記録する映像信号及び音声信号の記録再生装置が提案されている。

【0020】次にこの先に提案されている映像信号及び音声信号の記録再生装置につき図3～図7を参照して説明する。この先に提案されている映像信号及び音声信号の記録再生装置ではATRACにより圧縮される音声データのサンプリング周波数をDVD等で用いられている48.0kHzとする。48.0kHzのサンプリング周波数では、1GOPの中に24,024個の整数個のサンプリング点が存在することになる。それによって、映像信号と音声信号との同期をとることを可能とするものである。なお、サンプリング周波数としては、48.0kHzに限らず他の周波数、例えば96.0kHzを使用しても良い。

【0021】この映像信号及び音声信号の記録再生装置の例の全体ブロック図を図3に示す。まず、記録側において、21で示すビデオカメラから音声信号及び映像信号が出力される。ビデオカメラ21から出力される音声信号はMIC. OUTとして、Lチャンネルの音声信号がA/D変換器22へ供給され、Rチャンネルの音声信号がA/D変換器23へ供給される。A/D変換器22及び23は、クロック発生器(CLK GEN)25か

ら供給された 48.0 kHz の周波数のサンプリングクロックによって、デジタル化を行う。

【0022】A/D変換器22及び23からの音声データは、メモリ26へ供給される。メモリ26は、映像データと音声データのタイミングを調整するためのものである。CPU32の制御によって、メモリ26から音声データがATRACエンコーダ28へ供給される。音声データは、例えば16ビットからなるリニアPCM (Pulse Code Modulation) データであり、時分割されたLチャンネルとRチャンネルの音声信号である。ATRACエンコーダ28では、CPU32からのコントロール信号に基づいて、後述するようにGOPに対応する音声データが47個のサウンドグループに分けられる。サウンドグループに分けられた音声信号は、マルチプレクサ(MUX)30へ供給される。

【0023】映像信号は、ビデオカメラ21からCCD OUTとして、A/D変換器24へ供給される。A/D変換器24には、クロック発生器25から13.5 kHzの周波数のサンプリングクロックが供給され、映像信号がデジタル化される。デジタル化された映像データは、カメラプロセス回路27へ供給される。カメラプロセス回路27では、γ補正及びアパーチャ補正等のカメラ信号処理が行われ、輝度データY及び色差データCが生成される。この色差データCは、時分割された色差データCb (青の色差データ) 及び色差データCr (赤の色差データ) からなる。輝度データY及び色差データCは、それぞれ8ビットのデータとしてMPEGエンコーダ29へ供給される。MPEGエンコーダ29では、CPU32からのGOPを指示する信号等のコントロール信号に基づいて、輝度データY及び色差データCがMPEG2で符号化される。符号化された映像データは、マルチプレクサ30へ供給される。

【0024】マルチプレクサ30は、符号化された音声データと符号化された映像データとCPU32からのヘッダ情報とが供給され、ヘッダ情報の付加、音声信号の例外処理を行う。マルチプレクサ30からは、図6に示すように映像信号と音声信号とヘッダ情報が合成された記録データが出力される。記録データは、チャンネルコーディング回路31へ供給される。チャンネルコーディング回路31では、(1, 7) RLLの変調、スクランブル及びエラー訂正のエンコーディング等の処理が行なわれる。これらの処理が施されたデータは、ビットストリームで記録媒体33へ記録される。この記録媒体33は、磁気又は光テープ、MO (光磁気) ディスク等の記録可能なディスク、メモリカード等の半導体メモリである。

【0025】次に再生側において、記録媒体33からビットストリームで読み出されたデータは、チャンネルデコーディング回路41へ供給される。チャンネルデコーディング回路41では、(1, 7) RLLの復調、ディ

スクランブル及びエラー訂正等の処理が行なわれる。これらの処理が行なわれたデータは、ディマルチプレクサ42へ供給される。ディマルチプレクサ42では、供給されたデータからヘッダ情報、映像データ及び音声データが分離され、音声データはATRACデコーダ43へ供給され、映像データはMPEGデコーダ44へ供給される。また、ディマルチプレクサ42で分離されるヘッダ情報は、CPU32へ供給される。CPU32では、供給されたヘッダ情報から音声信号の例外処理の位置等の復号に必要な情報が検出され、その情報はコントロール信号としてATRACデコーダ43及びMPEGデコーダ44へ供給される。

【0026】ATRACデコーダ43では、CPU32からのコントロール信号に基づいて、ATRACで符号化された音声データが復号される。復号された音声データは、時分割されたLチャンネルとRチャンネルの音声信号であり、メモリ45へ供給される。メモリ45は、CPU32の制御によって、映像データと音声データのタイミングを調整するものである。メモリ45から読み出された音声データは、デジタルフィルタ47へ供給されるデジタルフィルタ47は、後段のD/A変換器48及び49に対するプリフィルタであり、Lチャンネルの音声データは、D/A変換器48へ供給され、Rチャンネルの音声データは、D/A変換器49へ供給される。アナログ化された音声信号は、それぞれ出力端子50及び51を介してTVモニタ53の音声信号入力端子へ供給される。

【0027】ディマルチプレクサ42から映像データが供給されたMPEGデコーダ44では、CPU32からのGOPを指示する信号等のコントロール信号に基づいて、MPEG2の復号を行う。復号された輝度データY及び色差データCは、それぞれ8ビットのデータとしてビデオエンコーダ46へ供給される。復号された色差データCは、時分割された色差データCr及び色差データCbである。ビデオエンコーダ46では、供給された輝度データYと色差データCからNTSC信号又はRGB信号が復号される。ビデオエンコーダ46には、D/A変換器が含まれ、ビデオエンコーダ46からアナログのNTSC信号又はRGB信号が出力される。NTSC信号又はRGB信号は、出力端子52を介してTVモニタ53の映像信号入力端子へ供給される。

【0028】ここで、従来のMPEG2によって圧縮された映像データとATRACによって圧縮された音声データと同期して記録する方法の一例を具体的に説明する。音声信号のサンプリング周波数は、GOP周波数(2000/1001Hz)の整数倍の周波数が設定される。具体的には、GOP周波数とサンプリング周波数の最小公倍数が小さく、一般的に用いられている48.0 kHzを音声信号のサンプリング周波数とする。このサンプリング周波数から求められる1GOP中の音声デ

ータのサンプル数は、24024個である。ATRACエンコーダは、8通りのモードが定義されており、この例では、一例としてコーディングモード：ATRAC

(1), Level 1, Channel No. = 2について説明する。

【0029】このコーディングモードは、音声データの片チャンネルの512個のサンプル数を212バイトの圧縮データにエンコードするものである。この512個のサンプル数のATRACの処理単位をサウンドグループ(SG)と呼ぶことにする。以下、説明を容易とするために1チャンネルのみについて説明する。24024個の音声データのサンプル数を512個のサンプル単位に区切ると、46個のサウンドグループ+472個の音声データのサンプル数となる。従って、サウンドグループを47個分用意すれば、端数は生じるが、1GOP分の音声データを格納することができる。

【0030】この48.0kHzのサンプリング周波数でサンプリングされた1GOP分の音声データのサンプル数の24024個と1GOP分の音声データのサンプルとを比較すると、1GOP分の音声データのサンプル

数は、
 $47(SG) \times 512(サンプル) = 24064(サンプル)$

であるから、1GOPのサンプル数に対して40個のサンプルの端数が生じる。

【0031】この40個のサンプルを吸収するために、この例では、47個のサウンドグループのうち40個のサウンドグループについて、実際に使用される音声データのサンプル数を1個減らして511個として、符号化が実行される。このように、有効なサンプル数を1個減少させる処理が例外処理である。残りの7個のサウンドグループについては、音声データのサンプル数を512個として、符号化が実行される。この512個の音声データのサンプル数からなるサウンドグループは出来るだけ47個のサウンドグループ中に均等に配置させる。例えば、0, 7, 14, 21, 27, 34, 41番目に512個の音声データのサンプルからなるサウンドグループを配置させることによって、ほぼ均等に音声データを配置することができる。

【0032】 $40(SG) \times 511(サンプル) + 7(SG) \times 512(サンプル) = 24024(サンプル)$

となり、1GOP分の音声データのサンプル数となる。すなわち、1GOP(約0.5秒)を単位として、映像データおよび音声データが完結することになる。

【0033】ここで、音声データのサンプル数が511個のサウンドグループをATRACでエンコードするときの手法を簡単に説明する。上述したように、この例で使われているATRACのエンコードでは、1サウンドグループについて、512個の音声データのサンプル数

が必要である。しかしながら、1GOP分の音声データのサンプル数が24024個のため、サンプル数が511個のサウンドグループを40個作るようにした。このサンプル数が511個のサウンドグループをATRACでエンコードする場合、1つ前のサウンドグループの512番目のサンプルをダミーデータとして使用する。そのダミーデータ1個を加えて、サンプル数が512個となったサウンドグループがATRACでエンコードされる。

【0034】次に再生について、説明する。再生時、すなわちデコード時は、ATRACのデコードから出力される512個のサンプルのうち、後半511個のサンプルを実データとして使用すれば良い。最初の1サンプルは、エンコード時、ダミーデータという扱いだったので捨てる。サウンドグループの境目は、後述するように、オーバーラップ分があり、窓関数の掛ける処理がなされるので、波形の連続性は保たれる。よって、ダミーデータとなる1サンプルの処理は、従来に比べてオーバーラップのサンプル数が1サンプルだけ増えると考えれば良い。ただし、どのサウンドグループが511個のサンプルを出力するのか、または512個のサンプルを出力するのかは、エンコード時にヘッダ情報に記録しておく必要がある。

【0035】図6に示すように、記録媒体には、上述したようにGOPのヘッダ情報、15フレームからなる1GOP分(約0.5秒)の映像データ及び約0.5秒(1GOP分)に相当する47サウンドグループからなる音声データが最小単位として、扱われる。このとき、1サウンドグループは、511個又は512個のサンプルからなり、1GOPで24024個のサンプルが存在する。そして、15フレームからなる映像データは、MPEG2で圧縮処理が施され、47サウンドグループからなる音声データは、ATRACで圧縮処理が施される。また、例外処理によって有効な音声データのサンプル数が511個とされたサウンドグループの位置の情報がヘッダ情報に記憶される。このヘッダ情報と、圧縮が施された映像データと音声データとからなる記録データ(ビットストリーム)が(1, 7)RLLの変調され、記録媒体へ記録される。

【0036】この記録媒体に対して、1GOP単位で映像データと音声データとが記録される。まず、ヘッダ情報に続いて、1GOP分の圧縮映像データ、さらに続いて2チャンネル分の音声データが1GOPの期間で記録される。ヘッダ情報には、どのサウンドグループが511個又は512個のサンプル数かが記録されている。この記録される音声データは、 $(47(SG) \times 2(ch) = 94$ サウンドグループからなる。さらに、GOP単位で映像データと音声データの位相は、常に合致している。

【0037】そして、ユーザからGOP単位で再生箇所

の指定を受け取ると、指定場所からのビットストリームを復調した後、復調されたビットストリームは、映像データ及び音声データのデコーダへ入力される。映像データのデコーダは、MPEGデコーダ44であり、音声データのデコーダは、ATRACデコーダ43である。ATRACデコーダ43の出力は、一旦メモリ45に入力される。この音声データは、デコードがなされた後、メモリ45に記憶され、映像データとのタイミング合わせを行なって後段へ出力される。ATRACデコーダは43及びMPEGデコーダ44の出力は、GOP単位で完結するので、一度同期を取ればずれることはない。

【0038】また、ATRACデコーダ43から受け取る音声データは、サウンドグループ当り512個のサンプル数であったり、511個のサンプル数であったりするが、上述したように、この例外処理のなされたサウンドグループが1GOP毎にヘッダ情報に書き込まれているので、これを読み込むことによって、1GOP内のサンプル数が丁度24024個である音声データが得られる。

【0039】このように、この例では、1GOPで映像データ及び音声データが完結しているの、インサートやアフレコ等をGOP単位で行うことは全く問題なくできる。

【0040】図3中のATRACエンコーダ28の詳細なブロック図を図4に示す。メモリ26からの各サンプルが16ビットからなるPCMデータ（リニアPCM）が入力端子61を介してATRACエンコーダ28へ供給される。このPCMデータは、リングバッファ64へ供給される。また、入力端子62からサウンドグループのサンプル数情報がATRACエンコーダ28へ供給されると共に、出力端子63からの後段のマルチプレクサ30へ供給される。このサウンドグループのサンプル数情報は、アドレス・コントロール回路65へ供給される。

【0041】アドレス・コントロール回路65では、供給されたサウンドグループのサンプル数情報からこのサウンドグループが511個のサンプルからなるものか、512個のサンプルからなるものかが判断される。このアドレス・コントロール回路65でこのサウンドグループは、511個のサンプルからなるサウンドグループであると判断されると、リングバッファ64から読み出される開始アドレスが1つ前にずらされる。リングバッファ64は、図7Aに示すように、サウンドグループ3個分の容量からなり、入力PCMデータが記憶される。

【0042】乗算器66では、図7Aに示すように、リングバッファ64から読み出されたPCMデータと窓関数とが乗算される。この図7Aにおいて、実線部は、今回の1サウンドグループ分の処理であり、破線部は、次の1サウンドグループ分の処理である。その乗算結果は、帯域分割フィルタ2へ供給され、高域成分（H）と

中域成分（M）及び低域成分（L）へ分割される。高域成分（H）は、MDCT（Modified Discrete Cosine Transform）4へ供給され、その他の成分は、帯域分割フィルタ3へ供給される。帯域分割フィルタ3では、供給された成分が中域成分（M）と低域成分（L）へ分割される。分割された中域成分（M）は、MDCT5へ供給され、低域成分（L）は、MDCT6へ供給される。

【0043】MDCT4、5及び6では、供給された高域成分、中域成分及び低域成分に対して変形DCTの演算が施され、各周波数成分に分解される。MDCT4、5及び6の出力は、ビット割当計算回路8及び正規化／量子化回路7へ供給される。ビット割当計算回路8では、供給された各成分から計算によって、割り当てられる量子化ビット数が算出される。量子化ビット数を表わす情報は、正規化／量子化回路7へ供給されるとともに、出力端子10から出力され、後段のマルチプレクサ30へ供給される。

【0044】正規化／量子化回路7では、量子化ビット数の情報から量子化ビット数を決定し、その量子化ビット数を用いてMDCT4、5及び6から供給された高域成分、中域成分及び低域成分（AC成分）に対して量子化が実行される。さらに、量子化がなされた各AC成分に対して正規化が実行される。また、正規化／量子化回路7からは、処理後のAC成分がスペクトル情報として、DC成分とともに出力され、この出力が出力端子9を介して後段のマルチプレクサ30へ供給される。

【0045】ここで、音声データのサンプル数が511個のサウンドグループをATRACでエンコードする手法を具体的に説明する。この例で使われているATRACのエンコードは、1サウンドグループすなわち512個の音声データのサンプル数を単位として実行される。しかしながら、実際には、この512個のサンプル以外に前後をそれぞれ256個のサンプルを加えて、合計1024個のサンプルのデータを持ってきて、窓関数を掛けてATRACのエンコードは行なわれる。この窓関数は、図7に示すように、1024個の全サンプル数の両端を0でうめるような関数である。

【0046】当然、511個のサンプルを扱うときも1024個のサンプルをATRACエンコーダの入力として用いる。ただし、この場合の有効サンプル数は、1サンプル減り、258番目から768番目までの511個のサンプルとなる。すなわち、図7Bに示すように、 n 番目のサウンドグループをエンコードする場合、中心となるサウンドグループのサンプル数の512個と、そのサウンドグループの前後の256個のサンプルの合計1,024個がATRACエンコーダに入力される。

【0047】しかしながら、次の $n+1$ 番目のサウンドグループをエンコードする場合、サウンドグループのサンプル数は511個のため、 n 番目のサウンドグループの512番目を含めてサンプルを512個とし、その5

12個のサウンドグループの前後の256個のサンプルの合計1,024がATRACエンコーダに入力される。すなわち、サウンドグループのサンプルが511個の場合、1サンプルだけ重なって、処理されるが、この重なった部分は、n番目のサウンドグループの符号化の処理においては、本来の処理であり、今回(n+1番目)のサウンドグループの符号化の処理では、ダミーデータとして処理される。

【0048】図3中のATRACデコーダ43の詳細なブロック図を図5に示す。入力端子12から量子化ビット数の情報及び正規化係数の情報がATRACデコーダ43へ供給され、同様に入力端子11からスペクトル情報がATRACデコーダ43へ供給される。また、入力端子71からサウンドグループのサンプル数情報がATRACデコーダ43へ供給されると共に、出力端子73を介してCPU32へ供給される。

【0049】ATRACデコーダ43に供給された量子化ビット数の情報、正規化係数の情報及びスペクトル情報がスペクトル復元回路13へ供給される。スペクトル復元回路13では、供給されたスペクトル情報、正規化係数の情報及び量子化ビット数の情報からAC成分が生成される。生成されたAC成分の高域成分(H)は、IMDCT(逆MDCT)14へ供給される。同様にAC成分の中域成分(M)は、IMDCT15へ供給される。さらに、低域成分(L)のAC成分及びDC成分は、IMDCT16へ供給される。

【0050】IMDCT14,15及び16では、供給された各成分に対して逆MDCT処理が施される。IMDCT14で逆MDCT処理により得られた高域成分の信号は、帯域合成フィルタ17へ供給される。また、IMDCT15及び16で逆MDCT処理により得られた中域成分及び低域成分の信号は、帯域合成フィルタ18へ供給される。帯域合成フィルタ18では、中域成分の信号と低域成分の信号とが合成され、その合成信号は、帯域合成フィルタ17へ供給される。帯域合成フィルタ17では、高域成分と合成信号との合成が行なわれ、その合成信号は、乗算器74へ供給される。

【0051】乗算器74では、窓関数と合成信号との乗算が行われ、その乗算結果は、リングバッファ75に記録される。この窓関数は、図7に示すように、1024個の全サンプル数の両端を0でうめるような関数である。図7Aに示すように、今回の窓関数により得られた乗算結果と、次の窓関数により得られた乗算結果とのオーバーラップとをリングバッファ75に書き込む時は、前に書き込んだデータを読み出し、読み出したデータと現データとを加算する。このように合成された窓関数は、リングバッファ75に記録される。このリングバッファ75は、サウンドグループ3個分の容量からなる。アドレス・コントロール回路76では、入力端子71から供給されるサウンドグループのサンプル数情報か

らリングバッファ75に記録されたサウンドグループが511個のサンプルからなるものか、512個のサンプルからなるものかが判断される。このアドレス・コントロール回路76でこのサウンドグループは、511個のサンプルからなるサウンドグループであると判断されると、リングバッファ75の開始アドレスが1つ前にずらされる。リングバッファ75から16ビットからなるPCMデータが読み出されるときに、ダミーデータの1サンプル分を取り除き、出力端子72から後段のメモリ45へ供給される。

【0052】斯る図3～図7に示す如き映像信号及び音声信号の記録再生装置によればMPEGで圧縮された映像データとATRACで圧縮された音声データとを同期して記録媒体に記録することができる。

【0053】またこの例によればGOP単位で記録し、再生を行うことができ、このときどのGOPから再生を開始しても映像と音声との同期(リップシンク)が守られている。

【0054】

20 【発明が解決しようとする課題】斯る従来の映像信号及び音声信号の記録再生装置においてはATRACエンコーダ28及びATRACデコーダ43内のリングバッファ64及び75を外側のCPU32からコントロールすることによって1サウンドグループ当りのサンプル数を適宜511個又は512個にして同期関係を保つ如くしている。

【0055】然しながら、これは偶数サンプル及び奇数サンプルのサウンドグループが並び合っていることであり、窓関数を掛けて合成する際に原理的にわずかながら振幅の変動を伴う不都合があった。この振幅変動はシミュレーションの結果、約0.2dBであった。

【0056】また上述従来例ではリングバッファ64及び75のポインタをコントロールする必要があるため現在MDに使用されているATRACエンコーダ及びATRACデコーダがそのまま使用できない(新規にLSIを作る必要がある。)不都合があった。

【0057】本発明は斯る点に鑑み上述振幅の変動を回避すると共に現在使用されているATRACエンコーダ及びATRACデコーダを使用できるようにすることを目的とする。

【0058】

40 【課題を解決するための手段】本発明映像信号及び音声信号の記録装置は映像データをMPEGで符号化し、圧縮映像データを生成する映像符号化手段と、この圧縮映像データのGOP単位で一定となるサンプル数が含まれるように音声信号のサンプリング周波数を設定し、このサンプリング周波数で、この音声信号をデジタル化するA/D変換手段と、このA/D変換手段からの音声データをATRACで符号化すると共にこの圧縮映像データのI GOPに対し一定サウンドグループ数の圧縮音声デ

ータを生成する音声符号化手段と、この音声符号化手段に供給する音声データのサンプル数を1サウンドグループ当たり一定とするダミーデータを付加するダミーデータ付加手段と、この圧縮映像データとこの圧縮音声データとをこのGOP単位で記録する記録手段とを有するものである。

【0059】斯る本発明によれば映像信号をMPEGで圧縮を施し、音声信号をATRACで圧縮を施した後、1GOPを単位として、同期が取られ記録媒体に記録することができる。この場合、本発明においては、ATRACで符号化する音声符号化手段に供給する音声データのサンプル数を1サウンドグループ当たり一定例えば512個とするダミーデータを付加するようにしたので、各サウンドグループのサンプル数は例えば512個と一定であり、窓関数を掛けて合成する際に振幅の変動を伴うことがない。

【0060】また本発明によれば、各サウンドグループのサンプル数が例えば512個と一定なので、リングバッファのポインタをコントロールする必要がなく、従来のMDに使用されていたATRACエンコーダ及びATRACデコーダがそのまま使用できる。

【0061】

【発明の実施の形態】以下図面を参照して本発明映像信号及び音声信号の記録装置を映像信号及び音声信号の記録再生装置に適用した例につき説明しよう。本例による全体ブロック図は前述図3と同様であり、本例においては図3のATRACエンコーダ28を図1に示す如く構成すると共にATRACデコーダ43を図2に示す如く構成したものである。

【0062】本例につき説明するに図3の全体ブロック図の説明は前述したので、これを省略する。

【0063】ここで、本例によるMPEG2によって圧縮された映像データとATRACによって圧縮された音声データとを同期して記録する例につき具体的に説明する。

【0064】音声信号のサンプリング周波数は、GOP周波数(2000/1001Hz)の整数倍の周波数が設定される。具体的には、GOP周波数とサンプリング周波数の最小公倍数が小さく、一般的に用いられている48.0kHzを音声信号のサンプリング周波数とする。このサンプリング周波数から求められる1GOP中の音声データのサンプル数は、24024個である。ATRACエンコーダは、8通りのモードが定義されており、この例では、一例としてコーディングモード:ATRAC(1)、Level1、ChannelNo.=2について説明する。

【0065】このコーディングモードは、音声データの片チャンネルの512個のサンプル数を212バイトの圧縮データにエンコードするものである。この512個のサンプル数のATRACの処理単位をサウンドグルー

プ(SG)と呼ぶことにする。以下、説明を容易とするために1チャンネルのみについて説明する。24024個の音声データのサンプル数を512個のサンプル単位に区切ると、46個のサウンドグループ+472個の音声データのサンプル数となる。従って、サウンドグループを47個分用意すれば、端数は生じるが、1GOP分の音声データを格納することができる。

【0066】この48.0kHzのサンプリング周波数でサンプリングされた1GOP分の音声データのサンプル数の24024個と1GOP分の音声データのサンプルとを比較すると、1GOP分の音声データのサンプル数は、

$$47(SG) \times 512(\text{サンプル}) = 24064(\text{サンプル})$$

であるから、1GOPのサンプル数(24024個)に対して40個のサンプリングの端数が生じる。

【0067】現在使用されているATRACエンコーダにおけるサウンドグループのサンプル数が512個なので、本例においてはATRACエンコーダ28として現在使用されているATRACエンコーダを使用し、このATRACエンコーダ28に1サウンドグループ当たり512個の一定のサンプル数の音声データを供給する如くする。

【0068】この場合前述の如く1GOPの音声データのサンプル数は24024個なので40個のデータが不足するが、本例においてはこのATRACエンコーダ28の前段に図1に示す如くダミーデータ付加手段を設け、適当なタイミングでこの40個のダミーデータを音声データには付加する如くする。

【0069】この場合512個の音声データのサンプル数からなるサウンドグループは出来るだけ47個のサウンドグループ中に均等に配置するを可とするので、例えば0, 7, 14, 27, 34, 41番目にこの512個の音声データのサンプルからなるサウンドグループを配置し、その他の40個のサウンドグループに1個づつダミーデータを付加する如くする。このサンプル数が512個の一定の単位のサウンドグループであり、且つ1GOPに対し47サウンドグループの音声データがATRACエンコーダ28に供給され、エンコードされる。

【0070】この音声データにダミーデータを付加した情報即ち何番目のサウンドグループのどの部分にダミーデータを付加したかの情報をヘッダ等別エリアに記録する如くする。

【0071】次に再生について説明する。本例のATRACデコーダ43においては、ATRACエンコーダ28で1サウンドグループを音声データのサンプル数を一定の512個としたので、このATRACデコーダ側でも同様の処理を行う必要がある(この場合現在存在するATRACデコーダをそのまま使用できる。))。

【0072】従ってこのATRACデコーダ43の出力

側には 1 GOP 当たり

$47 \text{ (SG)} \times 512 \text{ (サンプル)} = 24064 \text{ (サンプル)}$

のサンプル数が得られる。この 24064 サンプル数よりダミーデータ付加情報に従ってダミーデータを除去すれば、1 GOP 当たりの有効データの 24024 サンプル数の音声データを得ることができる。

【0073】この場合ダミーデータを取り除くには図 3 のメモリ 45 より音声データを読み出すときに CPU 32 よりの制御信号により取り除くようにすれば良い。

【0074】図 6 に示すように、記録媒体には、上述したように GOP のヘッダ情報、15 フレームからなる 1 GOP 分 (約 0.5 秒) の映像データ及び約 0.5 秒 (1 GOP 分) に相当する 47 サウンドグループからなる音声データが最小単位として、扱われる。このとき、1 サウンドグループは、512 個のサンプルからなり、1 GOP で 24064 個のサンプルが存在する。そして、15 フレームからなる映像データは、MPEG 2 で圧縮処理が施され、47 サウンドグループからなる音声データは、ATRAC で圧縮処理が施される。また、ダミーデータが付加されたサウンドグループの位置の情報がヘッダ情報に記憶される。このヘッダ情報と、圧縮が施された映像データと音声データとからなる記録データ (ビットストリーム) が (1, 7) RLL の変調され、記録媒体へ記録される。

【0075】この記録媒体に対して、1 GOP 単位で映像データと音声データとが記録される。まず、ヘッダ情報に続いて、1 GOP 分の圧縮映像データ、さらに続いて 2 チャンネル分の音声データが 1 GOP の期間で記録される。ヘッダ情報には、どのサウンドグループのどの位置にダミーデータが付加されたかが記録されている。この記録される音声データは、 $(47 \text{ (SG)} \times 2 \text{ (ch)} =) 94 \text{ サウンドグループ}$ からなる。さらに、GOP 単位で映像データと音声データの位相は、常に合致している。

【0076】そして、ユーザから GOP 単位で再生箇所の指定を受け取ると、指定場所からのビットストリームを復調した後、復調されたビットストリームは、映像データ及び音声データのデコーダへ入力される。映像データのデコーダは、MPEG デコーダ 44 であり、音声データのデコーダは、ATRAC デコーダ 43 である。ATRAC デコーダ 43 の出力は、一旦メモリ 45 に入力される。この音声データは、デコードがなされた後、メモリ 45 に記憶され、映像データとのタイミング合わせを行なうと共にダミーデータが除去されて後段へ出力される。ATRAC デコーダ 43 及び MPEG デコーダ 44 の出力は、GOP 単位で完結するので、一度同期を取ればずれることはない。

【0077】また、ATRAC デコーダ 43 から受け取る音声データは、サウンドグループ当たり 512 個のサ

ンプル数であるが、上述したように、メモリ 45 より読み出すときに 40 個のダミーデータが除去され 1 GOP 内のサンプル数が丁度 24024 個である音声データが得られる。

【0078】このように、この例では、1 GOP で映像データ及び音声データが完結しているため、インサートやアフレコ等を GOP 単位で行うことは全く問題なくできる。

【0079】図 3 中の ATRAC エンコーダ 28 の詳細なブロック図を図 1 に示す。メモリ 26 からの各サンプルが 16 ビットからなる PCM データ (リニア PCM) を入力端子 61 に供給する。

【0080】本例においては、この入力端子 61 に供給される 16 ビットからなる CPM データを切換スイッチ 80 の一方の固定接点 80a に供給すると共に 1 サンプル分の遅延素子 81 を介して他方の固定接点 80b に供給し、この切換スイッチ 80 の可動接点 80c をダミーデータ付加情報入力端子 82 に供給される CPU 32 よりのダミーデータ付加情報により切換える如くし、この可動接点 80c に 1 サウンドグループ当たり 512 個のサンプル数の PCM データを得る如くする。

【0081】この可動接点 80c に得られる PCM データを、リングバッファ 83 へ供給する。また、ダミーデータ付加情報入力端子 82 からのサウンドグループダミーデータ付加情報を出力端子 85 から後段のマルチプレクサ 30 へ供給する。

【0082】このリングバッファ 83 よりの 1 サウンドグループ当たり 512 個のサンプル数の PCM データを乗算器 84 を介して帯域分割フィルタ 2 に供給する。

【0083】乗算器 84 では、リングバッファ 83 から読み出された PCM データと窓関数とが乗算される。その乗算結果は、帯域分割フィルタ 2 へ供給され、高域成分 (H) と中域成分 (M) 及び低域成分 (L) へ分割される。高域成分 (H) は、MDCT (Modified Discrete Cosine Transform) 4 へ供給され、その他の成分は、帯域分割フィルタ 3 へ供給される。帯域分割フィルタ 3 では、供給された成分が中域成分 (M) と低域成分 (L) へ分割される。分割された中域成分 (M) は、MDCT 5 へ供給され、低域成分 (L) は、MDCT 6 へ供給される。

【0084】MDCT 4、5 及び 6 では、供給された高域成分、中域成分及び低域成分に対して変形 DCT の演算が施され、各周波数成分に分解される。MDCT 4、5 及び 6 の出力は、ビット割当計算回路 8 及び正規化／量子化回路 7 へ供給される。ビット割当計算回路 8 では、供給された各成分から計算によって、割り当てられる量子化ビット数が算出される。量子化ビット数を表わす情報は、正規化／量子化回路 7 へ供給されるとともに、出力端子 10 から出力され、後段のマルチプレクサ 30 へ供給される。

【0085】正規化／量子化回路7では、量子化ビット数の情報から量子化ビット数を決定し、その量子化ビット数を用いてMDCT4、5及び6からの供給された高域成分、中域成分及び低域成分（AC成分）に対して量子化が実行される。さらに、量子化がなされた各AC成分に対して正規化が実行される。また、正規化／量子化回路7からは、処理後のAC成分がスペクトル情報として、DC成分とともに出力され、この出力が出力端子9を介して後段のマルチプレクサ30へ供給される。

【0086】この例で使われているATRACのエンコードは、1サウンドグループすなわち512個の音声データのサンプル数を単位として実行される。しかしながら、実際には、この512個のサンプル以外に前後をそれぞれ256個のサンプルを加えて、合計1024個のサンプルのデータを持ってきて、窓関数を掛けてATRACのエンコードは行なわれる。この窓関数は、1024個の全サンプル数の両端を0でうめるような関数である。

【0087】図3中のATRACデコーダ43の詳細なブロック図を図2に示す。入力端子12から量子化ビット数の情報及び正規化係数の情報がATRACデコーダ43へ供給され、同様に入力端子11からスペクトル情報がATRACデコーダ43へ供給される。また、入力端子86からサウンドグループのダミーデータ付加情報が出力端子87を介してCPU32へ供給される。

【0088】ATRACデコーダ43に供給された量子化ビット数の情報、正規化係数の情報及びスペクトル情報がスペクトル復元回路13へ供給される。スペクトル復元回路13では、供給されたスペクトル情報、正規化係数の情報及び量子化ビット数の情報からAC成分が生成される。生成されたAC成分の高域成分（H）は、IMDCT（逆MDCT）14へ供給される。同様にAC成分の中域成分（M）は、IMDCT15へ供給される。さらに、低域成分（L）のAC成分及びDC成分は、IMDCT16へ供給される。

【0089】IMDCT14、15及び16では、供給された各成分に対して逆MDCT処理が施される。IMDCT14で逆MDCT処理により得られた高域成分の信号は、帯域合成フィルタ17へ供給される。また、IMDCT15及び16で逆MDCT処理により得られた中域成分及び低域成分の信号は、帯域合成フィルタ18へ供給される。帯域合成フィルタ18では、中域成分の信号と低域成分の信号とが合成され、その合成信号は、帯域合成フィルタ17へ供給される。帯域合成フィルタ17では、高域成分と合成信号との合成が行なわれ、その合成信号は、乗算器88へ供給される。

【0090】乗算器88では、窓関数と合成信号との乗算が行われ、その乗算結果は、リングバッファ89に記録される。この窓関数は、1024個の全サンプル数の両端を0でうめるような関数である。

【0091】この出力端子90に得られるリングバッファ89よりの16ビットからなるPCMデータをメモリ45に供給すると共に出力端子87よりのダミーデータ付加情報をCPU32に供給し、このCPU32よりの制御信号に従ってこのメモリ45より映像データとのタイミングを合わせると共にダミーデータを除去した音声データを得る如くする。その他は先に図3で説明した通りの動作を行う。

【0092】本例によれば映像信号をMPEGで圧縮を施し、音声信号をATRACで圧縮を施した後、1GOPを単位として、同期が取られ、記録媒体に記録することができる。

【0093】この場合本例においてはATRACで符号化するATRACエンコード28に供給する音声データのサンプル数を1サウンドグループ当たり一定の512個とするダミーデータを付加するようにしたので、各サウンドグループのサンプル数は一定の512個であり窓関数を掛けて合成する際に振幅の変動を伴うことがない利益がある。

【0094】また本例によれば各サウンドグループのサンプル数が512個と一定なのでリングバッファ83及び89のポイントをコントロールする必要がなく、従来MDに使用されていたATRACエンコード及びATRACデコーダがそのまま使用できる利益がある。

【0095】また、本例においては上述の如く圧縮映像データ及び圧縮音声データはGOP単位で完結しているため、1GOPを最小単位として、映像の差し替えや音声のアフレコ等の編集を行うことができる。

【0096】更に、音声データは1GOPの中に47サウンドグループとして格納しているので、音声データの編集に限れば、他の場所に全く影響を与えることなく、サウンドグループ単位（約10msec）の編集ができる。但しこの場合有効サンプル数が511個か512個かはもとに記録された音声データのサンプル数に従う。

【0097】尚、上述例はダミーデータとして遅延素子81を設けて、前置保持する如く述べたが、このダミーデータとして前後のサンプルデータから平均値補間しても良い。この後段でATRACエンコードする都合上このダミーデータの付加によって余計な高域成分が生じないようにすることが望ましい。また、本発明は上述の例に限ることなく本発明の要旨を逸脱することなくその他種々の構成が採り得ることは勿論である。

【0098】

【発明の効果】本発明によれば映像信号をMPEGで圧縮を施し、音声信号をATRACで圧縮を施した後、1GOPを単位として、同期が取られ、記録媒体に記録することができる。

【0099】この場合本発明においてはATRACで符号化する音声符号化手段に供給する音声データのサンプル数を1サウンドグループ当たり一定の例えば512個

とするダミーデータを付加するようにしたので、各サウンドグループのサンプル数は一定の例えば512個であり窓関数を掛けて合成する際に振幅の変動を伴うことがない利益がある。

【0100】また本発明によれば各サウンドグループのサンプル数が例えば512個と一定なのでリングバッファのポイントをコントロールする必要がなく、従来MDに使用されたATRACエンコーダ及びATRACデコーダがそのまま使用できる利益がある。

【面の簡単な説明】

【図1】本発明映像信号及び音声信号の記録装置の例に使用されるATRACエンコーダの例を示すブロック図である。

【図2】図3例に使用されるATRACデコーダの例を示すブロック図である。

【図3】本発明を適用した映像信号及び音声信号の記録再生装置の例を示すブロック図である。

【図4】従来のATRACエンコーダの例を示すブロック図である。

【図5】従来のATRACデコーダの例を示すブロック図である。

【図6】MPEGで圧縮を行う映像データとATRACで圧縮を行う音声データとを説明するための線図である。

【図7】従来のATRACエンコーダの説明に供する線

図である。

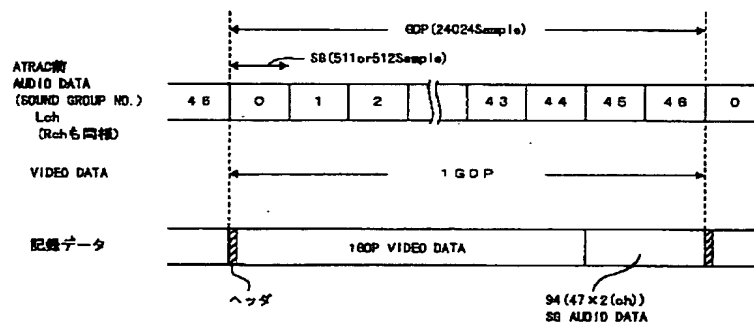
【図8】従来のATRACエンコーダの例を示すブロック図である。

【図9】従来のATRACデコーダの例を示すブロック図である。

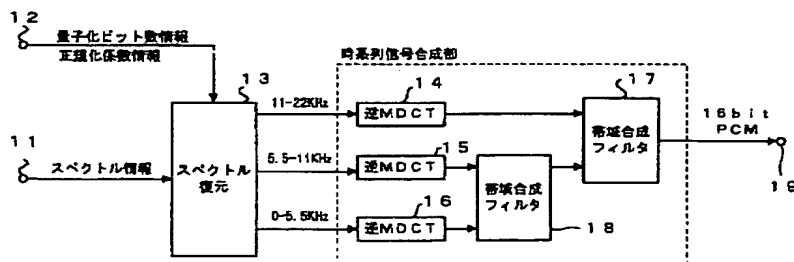
【符号の説明】

1…入力端子、2, 3…帯域分割フィルタ、4, 5, 6…MDCT回路、7…正規化/量子化回路、8…ビット割り当て算出回路、9, 10…出力端子、11, 12…入力端子、13…スペクトル復元回路、14, 15, 16…IMDCT回路、17, 18…帯域合成フィルタ、19…出力端子、21…ビデオカメラ、22, 23, 24…A/D変換器、25…クロック発生器、26, 45…メモリ、27…カメラプロセス回路、28…ATRACエンコーダ、29…MPEGエンコーダ、30…マルチプレкса、31…チャンネルコーディング回路、32…CPU、33…記録媒体、41…チャンネルデコーディング回路、42…逆マルチプレкса、43…ATRACデコーダ、44…MPEGデコーダ、46…ビデオエンコーダ、47…デジタルフィルタ、49…D/A変換器、53…TVモニタ、61, 62, 71, 82, 86…入力端子、63, 72, 73, 85, 87…出力端子、64, 75, 83, 89…リングバッファ、66, 74, 84, 88…乗算器、80…切換スイッチ、81…遅延素子

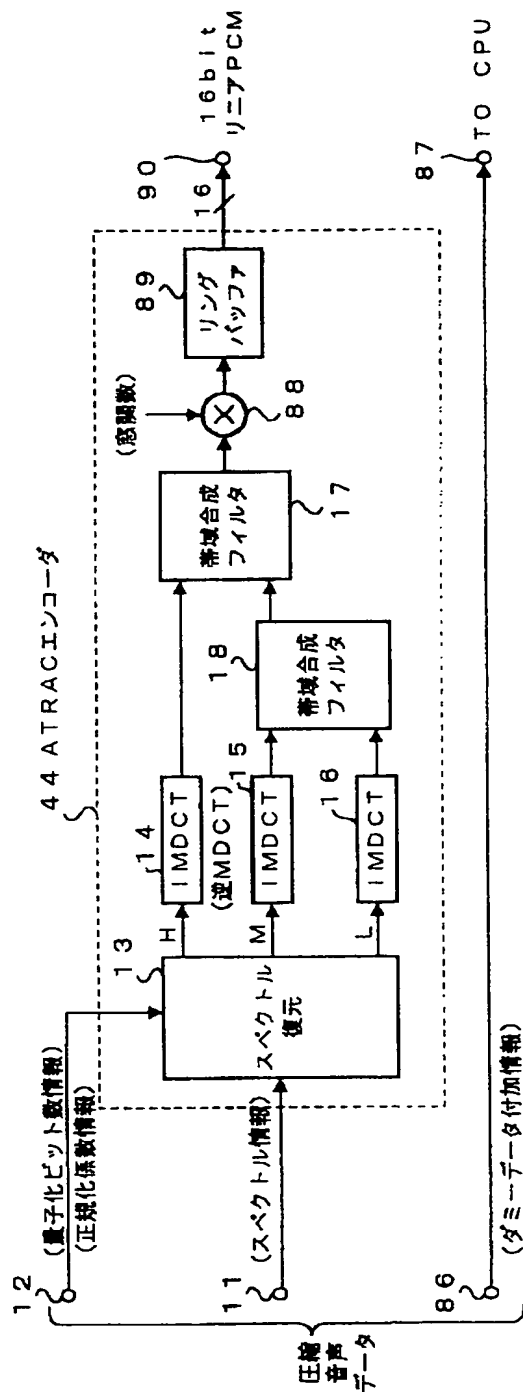
【図6】



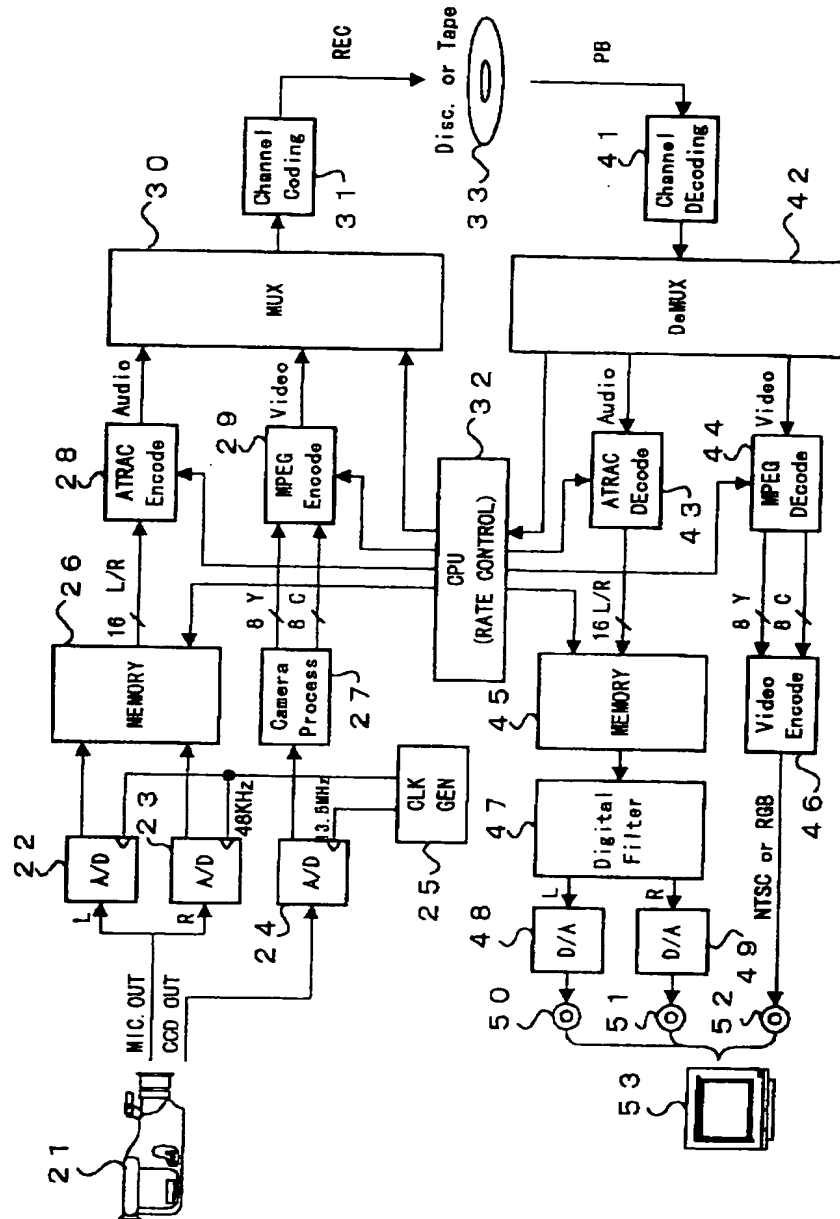
【図9】



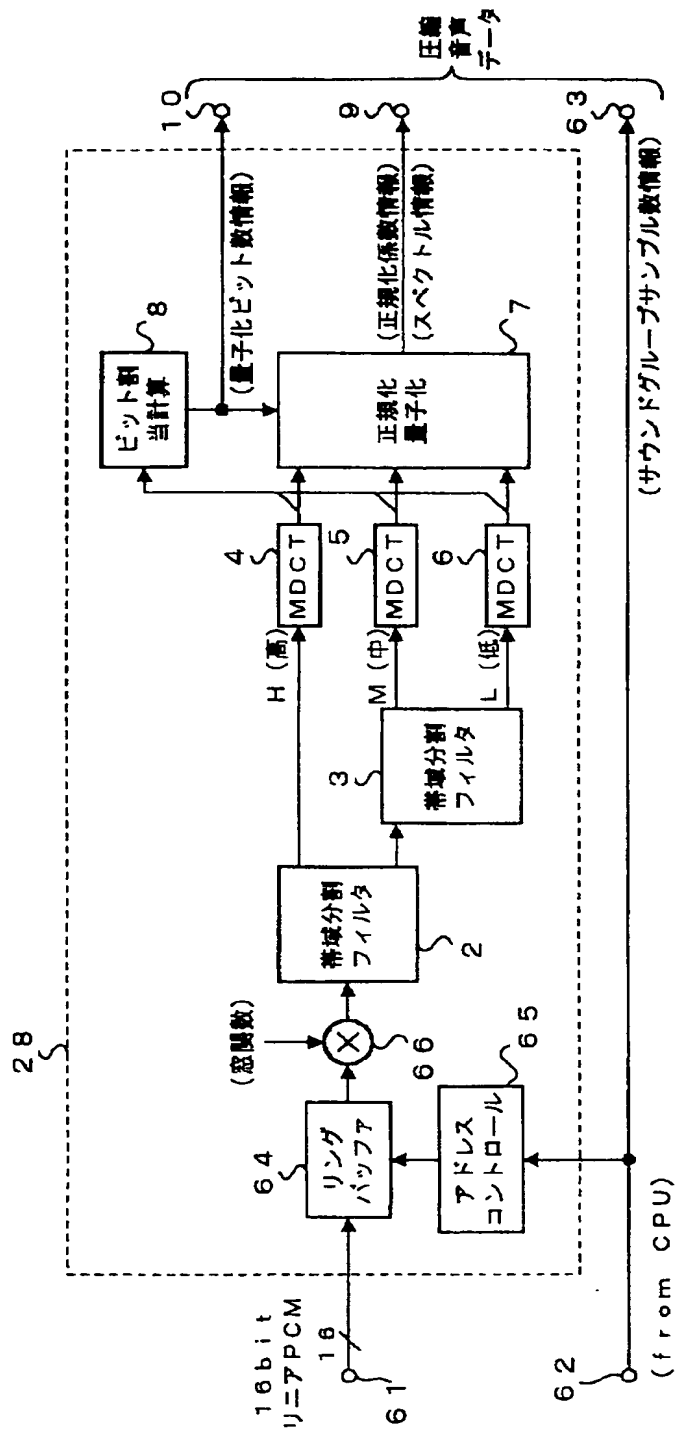
【図2】



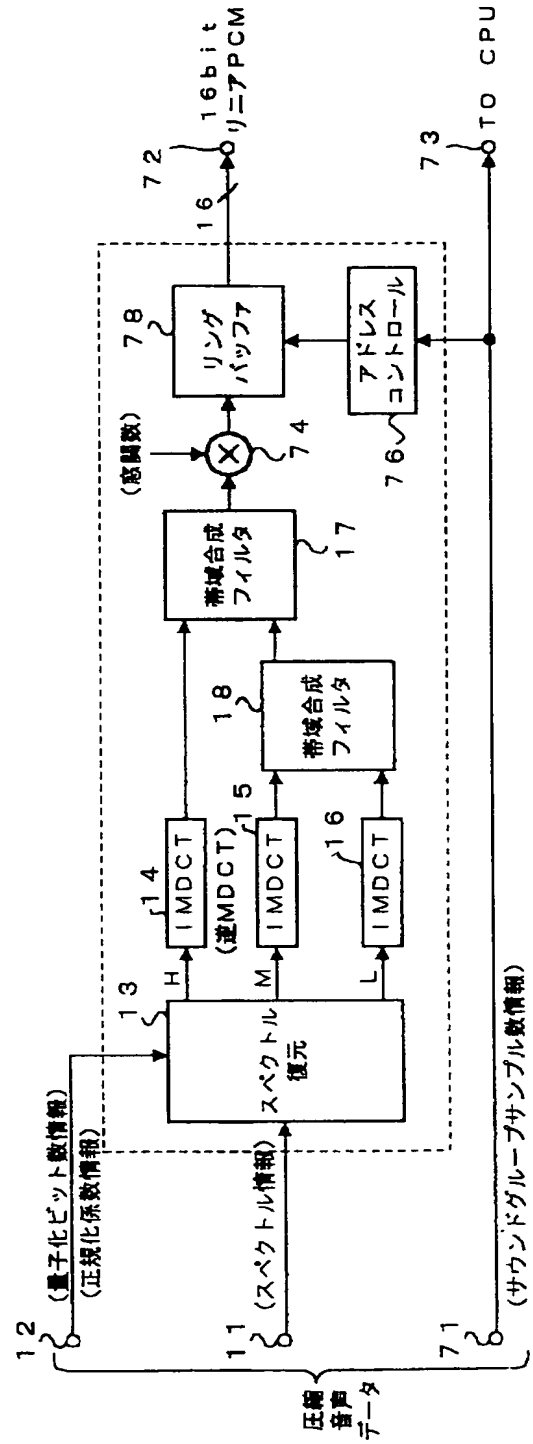
【図3】



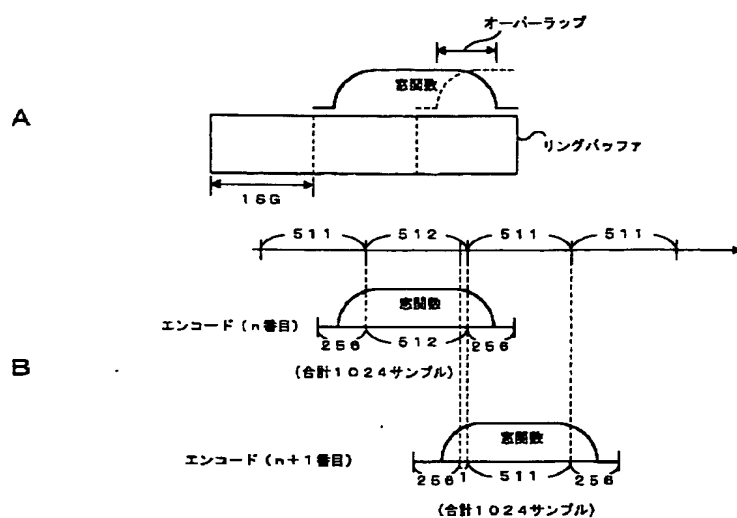
【図 4】



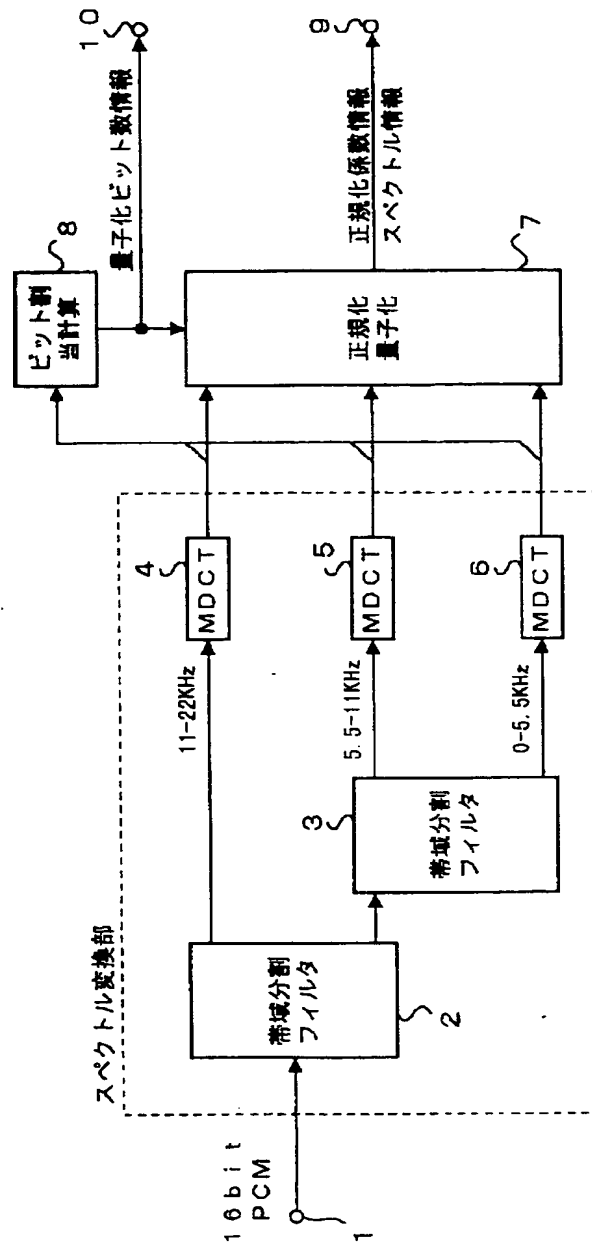
【図5】



【図 7】



【図 8】



フロントページの続き

(51) Int. Cl. ⁶

H 0 4 N 7/24

識別記号

F I

H 0 4 N 7/13

Z